AD-A083 033 VIRGINIA POLYTECHNIC INST AND STATE UNIV BLACKSBURG F/G 12/1
INEQUALITIES FOR JOINT DISTRIBUTIONS OF QUADRATIC FORMS.(U)
DEC 79 D R JENSEN DAAG29-78-G-0172
UNCLASSIFIED RR-Q-1 ARO-15194.2-M NL

1 OF 1
AD A083033

END
DATE FILMED
5-80
DTIC

TABLE OF CONTENTS

ITEM	PAGE
ABSTRACT	1
INTRODUCTION	1
THE MAIN RESULTS	2
APPLICATIONS	9
REFERENCES	13

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or special
A

INEQUALITIES FOR JOINT DISTRIBUTIONS OF QUADRATIC FORMS

D. R. Jensen

Virginia Polytechnic Institute
and State University

0. Abstract. Chebychev inequalities are given for joint central and noncentral distributions of k quadratic forms; these are sharpened when k=2 using the canonical correlations of Hotelling. Complementary inequalities are found as versions of Markov's inequality. Applications are noted in ballistics, in statistical quality control, in establishing consistency of Gauss-Markov estimates under dependence, and in constructing conservative joint confidence sets depending on the underlying distribution only through its low order moments.

1. Introduction. Chebychev inequalities are fundamental to statistical theory and practice, providing connections to laws of large numbers and conservative estimates for probabilities. Collections of quadratic forms arise in many branches of applied probability and statistics, for example as amplitudes of random signals in multichannel receivers and as meausres of discrepancy in parametric and nonparametric

---

AMS 1970 subject classifications. Primary 62H05; Secondary 62E15.

Key words and phrases. Multidimensional Chebychev and Markov inequalities, sharpened Chebychev bounds, weak laws of large numbers, ballistic probabilities, statistical quality control.

Acknowledgment. Supported in part by ARO Grant No. DAAG-29-78-G-0172 from the U.S. Army Research Office.

multiple testing procedures, where their central and noncentral distributions assume a prominent role. Numerous multidimensional Chebychev inequalities are available for probabilities of deviations from means. However, few of these apply to distributions of quadratic forms, then only in the central case, and even those results are scattered.

Here we develop Chebychev inequalities for joint central and noncentral distributions of k positive semidefinite quadratic forms, extending results of Wilks (1962); we sharpen these bounds for central distributions when k=2 using information supplied by Hotelling's (1936) canonical correlations, extending work of Berge (1937) and Lal (1955); we derive complementary inequalities as multidimensional versions of Markov's inequality; and we extend some of these findings to indefinite forms. We then apply these inequalities in a variety of problems in statistics and applied probability.

2. The Main Results. Let $\underset{\sim}{Y} = [Y_1,\ldots,Y_n]'$ be a random element in R^n having the mean $\underset{\sim}{\mu} = [\mu_1,\ldots,\mu_n]'$ and the nonsingular dispersion matrix $V(\underset{\sim}{Y}) = \underset{\sim}{\Sigma} = [\sigma_{ij}]$, and let $\{\underset{\sim}{B}_1,\ldots,\underset{\sim}{B}_k\}$ be any positive semidefinite matrices of order $(n\times n)$. A Chebychev inequality for the joint noncentral distribution of semidefinite quadratic forms in $\underset{\sim}{Y}$ is the following.

THEOREM 1. Let $\underset{\sim}{Y}$ be random having the finite mean $\underset{\sim}{\mu}$ and dispersion matrix $\underset{\sim}{\Sigma}$, and let $\{\underset{\sim}{B}_1,\ldots,\underset{\sim}{B}_k\}$ be positive semidefinite $(n\times n)$ matrices. For each $\underset{\sim}{\theta} \in R^n$ and each positive $\{\delta_1,\ldots,\delta_k\}$, we have

$$P((\underset{\sim}{Y}-\underset{\sim}{\theta})'\underset{\sim}{B}_1(\underset{\sim}{Y}-\underset{\sim}{\theta})\le\delta_1,\ldots,(\underset{\sim}{Y}-\underset{\sim}{\theta})'\underset{\sim}{B}_k(\underset{\sim}{Y}-\underset{\sim}{\theta})\le\delta_k) \ge 1-\left(\frac{\gamma_1}{\delta_1}+\ldots+\frac{\gamma_k}{\delta_k}\right)$$

where $\gamma_i = \mathrm{tr}\underset{\sim}{B}_i\underset{\sim}{\Sigma}+(\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}_i(\underset{\sim}{\mu}-\underset{\sim}{\theta})$ for $i=1,2,\ldots,k$.

Proof. Let $Q_i(\underset{\sim}{y}) = (\underset{\sim}{y}-\underset{\sim}{\theta})'\underset{\sim}{B}_i(\underset{\sim}{y}-\underset{\sim}{\theta})$ and $Q(\underset{\sim}{y}) = \delta_1^{-1}Q_1(\underset{\sim}{y})+\ldots+\delta_k^{-1}Q_k(\underset{\sim}{y})$ and, for $i=1,\ldots,k$, identify the sets $A_i = \{\underset{\sim}{y}\in R^n \mid Q_i(\underset{\sim}{y})\le\delta_i\}$ and $A = A_1\cap\ldots\cap A_k$. From the nonnegativity of $\{Q_1(\underset{\sim}{y}),\ldots,Q_k(\underset{\sim}{y})\}$ and the definition of A, we infer that $Q(\underset{\sim}{y}) \ge 0$ for all $\underset{\sim}{y}$ and $Q(\underset{\sim}{y}) > 1$ on the complement A^c of A. A standard computation yields

$$E[Q(\underset{\sim}{Y})] = \int_{A\cup A^c} Q(\underset{\sim}{y})dF(\underset{\sim}{y}) \ge \int_{A^c} Q(\underset{\sim}{y})dF(\underset{\sim}{y}) \ge \int_{A^c} dF(\underset{\sim}{y}) = 1-P(A)$$

with $F(\underset{\sim}{y})$ the cumulative distribution function of $\underset{\sim}{Y}$, and it remains to evaluate $E[Q(\underset{\sim}{y})]$. A routine argument gives $E[Q(\underset{\sim}{Y})] = \delta_1^{-1}[\mathrm{tr}\underset{\sim}{B}_1\underset{\sim}{\Sigma} + (\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}_1(\underset{\sim}{\mu}-\underset{\sim}{\theta})] + \ldots + \delta_k^{-1}[\mathrm{tr}\underset{\sim}{B}_k\underset{\sim}{\Sigma} + (\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}_k(\underset{\sim}{\mu}-\underset{\sim}{\theta})]$ and thus the theorem.

It is known for certain symmetric unimodal distributions on R^n that the measure of a convex symmetric set diminishes as the set is moved away from $\underset{\sim}{\mu}$ along a ray (cf. Anderson (1955)). We note here that a similar behavior with respect to ellipsoids is exhibited by any n-dimensional distribution having second-order moments, in the sense that the guaranteed probability diminishes as $(\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}(\underset{\sim}{\mu}-\underset{\sim}{\theta})$ increases. This is a consequence of the following corollary; if in addition $\underset{\sim}{\theta}=\underset{\sim}{\mu}$ and $\underset{\sim}{B}=\underset{\sim}{\Sigma}^{-1}$, the corollary yields $P((\underset{\sim}{Y}-\underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{Y}-\underset{\sim}{\mu})\le\delta) \ge 1-n/\delta$ as given in Wilks (1962; p. 92).

COROLLARY 1.1. If $\underset{\sim}{Y}$ be random with parameters $(\underset{\sim}{\mu},\underset{\sim}{\Sigma})$, then for each $\underset{\sim}{\theta}\in R^n$, each $\delta>0$, and each positive semidefinite matrix $\underset{\sim}{B}$,

$$P((\underset{\sim}{Y}-\underset{\sim}{\theta})'\underset{\sim}{B}(\underset{\sim}{Y}-\underset{\sim}{\theta})\le\delta) \ge 1-[\mathrm{tr}\underset{\sim}{B}\underset{\sim}{\Sigma}+(\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}(\underset{\sim}{\mu}-\underset{\sim}{\theta})]/\delta.$$

Many applications entail definite quadratic forms in subsets of the elements of $\underset{\sim}{Y}$. Partition $\underset{\sim}{Y} = [\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k']'$, $\underset{\sim}{\mu} = [\underset{\sim}{\mu}_1', \ldots, \underset{\sim}{\mu}_k']'$, $\underset{\sim}{\theta} = [\underset{\sim}{\theta}_1', \ldots, \underset{\sim}{\theta}_k']'$, and $\underset{\sim}{\Sigma} = [\underset{\sim}{\Sigma}_{ij}]$ conformably, with $\underset{\sim}{Y}_i$ of order $(m_i \times 1)$ and $m_1 + \ldots + m_k = n$, and choose block-diagonal matrices $\{\underset{\sim}{B}_1, \ldots, \underset{\sim}{B}_k\}$ of the types $\underset{\sim}{B}_1 = \text{Diag}(\underset{\sim}{E}_1, \underset{\sim}{0})$, $\underset{\sim}{B}_i = \text{Diag}(\underset{\sim}{0}, \underset{\sim}{E}_i, \underset{\sim}{0})$, and $\underset{\sim}{B}_k = \text{Diag}(\underset{\sim}{0}, \underset{\sim}{E}_k)$ with $\underset{\sim}{E}_i$ conforming to $\underset{\sim}{Y}_i$ for $i=1,\ldots,k$. Theorem 1 yields the following useful corollary, a special case of which was given in Wilks (1962; p. 274) when $k=n, m_1 = \ldots = m_k = 1$, and $\underset{\sim}{\theta} = \underset{\sim}{\mu}$.

<u>COROLLARY 1.2</u>. Let $\underset{\sim}{Y} = [\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k']'$ be random having the mean $\underset{\sim}{\mu} = [\underset{\sim}{\mu}_1', \ldots, \underset{\sim}{\mu}_k']'$ and the dispersion matrix $\underset{\sim}{\Sigma} = [\underset{\sim}{\Sigma}_{ij}]$, and let $\{\underset{\sim}{E}_i(m_i \times m_i);\ i=1,\ldots,k\}$ be positive semidefinite matrices. Then for each $\underset{\sim}{\theta} = [\underset{\sim}{\theta}_1', \ldots, \underset{\sim}{\theta}_k']' \in R^n$ and each positive $\{\delta_1, \ldots, \delta_k\}$, we have

$$P((\underset{\sim}{Y}_1 - \underset{\sim}{\theta}_1)'\underset{\sim}{E}_1(\underset{\sim}{Y}_1 - \underset{\sim}{\theta}_1) \le \delta_1, \ldots, (\underset{\sim}{Y}_k - \underset{\sim}{\theta}_k)'\underset{\sim}{E}_k(\underset{\sim}{Y}_k - \underset{\sim}{\theta}_k) \le \delta_k) \ge 1 - \left(\frac{\gamma_1}{\delta_1} + \ldots + \frac{\gamma_k}{\delta_k}\right)$$

where $\gamma_i = \text{tr}\underset{\sim}{E}_i\underset{\sim}{\Sigma}_{ii} + (\underset{\sim}{\mu}_i - \underset{\sim}{\theta}_i)'\underset{\sim}{E}_i(\underset{\sim}{\mu}_i - \underset{\sim}{\theta}_i)$ for $i=1,\ldots,k$.

Our inequalities thus far pertain to events of the type $A = \{\underset{\sim}{y} \in R^n | Q_1(\underset{\sim}{y}) \le \delta_1, \ldots, Q_k(\underset{\sim}{y}) \le \delta_k\}$. Complementary inequalities have to do with $P(Q_1(\underset{\sim}{Y}) \ge \delta_1, \ldots, Q_k(\underset{\sim}{Y}) \ge \delta_k)$. To establish such inequalities we proceed as follows. Let X be random with values in a linear space X; let $X_0 \in X$ be fixed; induce a partial order $x \prec y$ on X which is reflexive and transitive; and call a real valued function $g(\cdot)$ on X <u>monotone</u> if $x \prec y$ implies $g(x) \le g(y)$. The following general version of Markov's inequality is given in Jensen and Foutz (1978).

LEMMA 1. Let $X_0 \in X$ be fixed, and let $\mathcal{G}_0$ be the class of non-negative monotone functions on X depending on $P(\cdot)$ such that, for each $g \in \mathcal{G}_0$, $E[g(X)]$ is defined. Then

$$P(X \succ X_0) \leq \inf_{\mathcal{G}_0} E[g(X)]/g(X_0).$$

To complement Theorem 1 we identify X as R^n; we induce a partial order on R^n with respect to the positive semidefinite matrices $\{\underset{\sim}{B}_1, \ldots, \underset{\sim}{B}_k\}$ on stipulating that $\underset{\sim}{x}\ \underset{\sim}{y}$ if and only if $\underset{\sim}{x}'\underset{\sim}{B}_i\underset{\sim}{x} \leq \underset{\sim}{y}'\underset{\sim}{B}_i\underset{\sim}{y}$ for all $i=1,\ldots,k$; we choose for convenience the monotone function $g(\underset{\sim}{x}) = \underset{\sim}{x}'\underset{\sim}{B}\underset{\sim}{x}$ with $\underset{\sim}{B} = \underset{\sim}{B}_1+\ldots+\underset{\sim}{B}_k$; and we further assume that the several statements $\{\underset{\sim}{x}'\underset{\sim}{B}_1\underset{\sim}{x} \geq \delta_1, \ldots, \underset{\sim}{x}'\underset{\sim}{B}_k\underset{\sim}{x} \geq \delta_k\}$ are consistent. On letting $x_0 \in R^n$ be any point for which equality is achieved, i.e., $\{\underset{\sim}{x}_0'\underset{\sim}{B}_1\underset{\sim}{x}_0 = \delta_1, \ldots, \underset{\sim}{x}_0'\underset{\sim}{B}_k\underset{\sim}{x}_0 = \delta_k\}$, we apply Lemma 1 directly to establish the following inequality complementary to that of Theorem 1.

THEOREM 2. Let $\underset{\sim}{Y}$ be random having the finite mean $\underset{\sim}{\mu}$ and dispersion matrix $\underset{\sim}{\Sigma}$, and let $\{\underset{\sim}{B}_1, \ldots, \underset{\sim}{B}_k\}$ be positive semidefinite $(n\times n)$ matrices. For each $\underset{\sim}{\theta} \in R^n$ and each positive $\{\delta_1, \ldots, \delta_k\}$, we have

$$P((\underset{\sim}{Y}-\underset{\sim}{\theta})'\underset{\sim}{B}_1(\underset{\sim}{Y}-\underset{\sim}{\theta}) \geq \delta_1, \ldots, (\underset{\sim}{Y}-\underset{\sim}{\theta})'\underset{\sim}{B}_k(\underset{\sim}{Y}-\underset{\sim}{\theta}) \geq \delta_k) \leq \frac{\gamma_1+\ldots+\gamma_k}{\delta_1+\ldots+\delta_k}$$

where $\gamma_i = \mathrm{tr}\underset{\sim}{B}_i\underset{\sim}{\Sigma} + (\underset{\sim}{\mu}-\underset{\sim}{\theta})'\underset{\sim}{B}_i(\underset{\sim}{\mu}-\underset{\sim}{\theta})$ for $i=1,\ldots,k$.

Theorems 1 and 2 are basic to the present study, applying to joint distributions of positive semidefinite quadratic forms with bounds depending in part on the traces of specified matrices. We next inquire whether comparable inequalities may be found for indefinite quadratic forms, and whether the bounds given may be sharpened on using more

information about dispersion parameters. Both questions are answered affirmatively under the conditions set forth in the paragraphs following.

Suppose $\{B_1,\ldots,B_k\}$ are the matrices of indefinite quadratic forms and they commute. Then $\{B_1,\ldots,B_k\}$ are reducible to diagonal arrays by the same orthogonal matrix $Q(n\times n)$, and it thus suffices to consider a canonical form in which $\{B_1,\ldots,B_k\}$ themselves are diagonal and Y has the dispersion matrix $V(Y) = \Xi = Q'\Sigma Q$. Let $\{D_1^+,\ldots, D_k^+\}$ be the diagonal matrices obtained on replacing all negative elements of $\{B_1,\ldots,B_k\}$ by zeros in the canonical form of the problem; observe that $\{x'D_1^+x \le \delta_1,\ldots,x'D_k^+x \le \delta_k\}$ implies $\{x'B_1x \le \delta_1,\ldots, x'B_kx \le \delta_k\}$; and use earlier results as they apply to the positive semidefinite forms $\{x'D_1^+x,\ldots,x'D_k^+x\}$. An amended version of Theorems 1 and 2 is the following; versions of the corollaries follow similarly without difficulty.

<u>THEOREM 3</u>. Let Y be random having the finite mean μ and dispersion matrix Σ, and let $\{B_1,\ldots,B_k\}$ be indefinite commuting matrices reducible to the diagonal arrays $\{D_1,\ldots,D_k\}$. Then for each $\theta \in R^n$ and each positive $\{\delta_1,\ldots,\delta_k\}$, we have

$$P((Y-\theta)'B_1(Y-\theta)\le\delta_1,\ldots,(Y-\theta)'B_k(Y-\theta)\le\delta_k) \ge 1-\left(\frac{\lambda_1}{\delta_1}+\ldots+\frac{\lambda_k}{\delta_k}\right)$$

and

$$P((Y-\theta)'B_1(Y-\theta)\ge\delta_1,\ldots,(Y-\theta)'B_k(Y-\theta)\ge\delta_k) \le \frac{\lambda_1+\ldots+\lambda_k}{\delta_1+\ldots+\delta_k}$$

where $\lambda_i = \mathrm{tr}D_i^+\Xi + \eta'D_i^+\eta$, D_i^+ is the positive part of the diagonal matrix $D_i = Q'B_iQ$ with other elements replaced by zeros, $\eta = Q'(\mu-\theta)$, and $\Xi = Q'\Sigma Q$.

In order to sharpen Chebychev's inequality through the use of further information about dispersion parameters, we consider the special case of the problem treated in Theorem 1 with $k=2$, $\underset{\sim}{\theta}=\underset{\sim}{\mu}$, $\underset{\sim}{B}_1=\underset{\sim}{\Sigma}_{11}^{-1}$, and $\underset{\sim}{B}_2=\underset{\sim}{\Sigma}_{22}^{-1}$. This case leads naturally to a reduced form in which the canonical correlation parameters of Hotelling (1936) assume a prominent role. Our main result is the following.

THEOREM 4. Let $\underset{\sim}{Y} = [\underset{\sim}{Y}_1', \underset{\sim}{Y}_2']'$ be random of order $r + s = n$, with $r \le s$, having the finite mean $\underset{\sim}{\mu} = [\underset{\sim}{\mu}_1', \underset{\sim}{\mu}_2']'$, the dispersion matrix $\underset{\sim}{\Sigma} = [\underset{\sim}{\Sigma}_{ij}]$, and the canonical correlations $\{\rho_1,\ldots,\rho_r\}$. For each positive $\{\delta_1,\delta_2\}$, we have

$$P((\underset{\sim}{Y}_1-\underset{\sim}{\mu}_1)'\underset{\sim}{\Sigma}_{11}^{-1}(\underset{\sim}{Y}_1-\underset{\sim}{\mu}_1)\le\delta_1,(\underset{\sim}{Y}_2-\underset{\sim}{\mu}_2)'\underset{\sim}{\Sigma}_{22}^{-1}(\underset{\sim}{Y}_2-\underset{\sim}{\mu}_2)\le\delta_2) \ge 1-R(\delta_1,\delta_2;\underset{\sim}{\rho})$$

where $\underset{\sim}{\rho} = [\rho_1,\ldots,\rho_r]'$ and

$$R(\delta_1,\delta_2;\underset{\sim}{\rho}) = \frac{(s-r)}{\delta_2} + \sum_{i=1}^{r}\{(\delta_1+\delta_2)+[(\delta_1+\delta_2)^2-4\rho_i^2\delta_1\delta_2]^{\frac{1}{2}}\}/2\delta_1\delta_2.$$

Proof. Choose symmetric square roots and let $\underset{\sim}{Z} = [\underset{\sim}{Z}_1', \underset{\sim}{Z}_2']'$ such that $\underset{\sim}{Z}_1 = \underset{\sim}{\Sigma}_{11}^{-\frac{1}{2}}(\underset{\sim}{Y}_1-\underset{\sim}{\mu}_1)$ and $\underset{\sim}{Z}_2 = \underset{\sim}{\Sigma}_{22}^{-\frac{1}{2}}(\underset{\sim}{Y}_2-\underset{\sim}{\mu}_2)$; observe that $E(\underset{\sim}{Z}) = \underset{\sim}{0}$, $E(\underset{\sim}{Z}_1\underset{\sim}{Z}_1') = \underset{\sim}{I}_r$, $E(\underset{\sim}{Z}_1\underset{\sim}{Z}_2') = \underset{\sim}{R}_{12} = \underset{\sim}{\Sigma}_{11}^{-\frac{1}{2}}\underset{\sim}{\Sigma}_{12}\underset{\sim}{\Sigma}_{22}^{-\frac{1}{2}}$, and $E(\underset{\sim}{Z}_2\underset{\sim}{Z}_2') = \underset{\sim}{I}_s$; and, by invariance of $\underset{\sim}{Z}_1'\underset{\sim}{Z}_1$ and $\underset{\sim}{Z}_2'\underset{\sim}{Z}_2$ under orthogonal transformation, choose orthogonal matrices $\underset{\sim}{P}(r\times r)$ and $\underset{\sim}{Q}(s\times s)$ so as to achieve the singular decomposition $\underset{\sim}{P}\underset{\sim}{R}_{12}\underset{\sim}{Q}' = \underset{\sim}{D} = [\mathrm{Diag}(\rho_1,\ldots,\rho_r),\underset{\sim}{0}]$, where $\{\rho_1,\ldots,\rho_r\}$ are the canonical correlation parameters of Hotelling (1936). In canonical form with $\underset{\sim}{U}_1 = \underset{\sim}{P}\underset{\sim}{Z}_1/\delta_1^{\frac{1}{2}}$ and $\underset{\sim}{U}_2 = \underset{\sim}{Q}\underset{\sim}{Z}_2/\delta_2^{\frac{1}{2}}$, it suffices to demonstrate that $P(A) \ge 1-R(\delta_1,\delta_2;\underset{\sim}{\rho})$, where $A = \{\underset{\sim}{U} | \underset{\sim}{U}_1'\underset{\sim}{U}_1\le 1, \underset{\sim}{U}_2'\underset{\sim}{U}_2\le 1\}$ and A^c is its complement. To this end consider $g(\underset{\sim}{u};\underset{\sim}{t}) = \underset{\sim}{u}'[\underset{\sim}{G}(\underset{\sim}{t})]^{-1}\underset{\sim}{u}$ as a

function of the adjustable parameters $\underset{\sim}{t} = [t_1,\ldots,t_r]'$, where $\underset{\sim}{G}(\underset{\sim}{t}) = [\underset{\sim}{G}_{ij}(\underset{\sim}{t})]$ is a partitioned matrix with $\underset{\sim}{G}_{11}(\underset{\sim}{t}) = \underset{\sim}{I}_r$, $\underset{\sim}{G}_{22}(\underset{\sim}{t}) = \underset{\sim}{I}_s$, and $\underset{\sim}{G}_{12}(\underset{\sim}{t}) = \underset{\sim}{G}_{21}'(\underset{\sim}{t}) = [\mathrm{Diag}(t_1,\ldots,t_r),\underset{\sim}{0}] = \underset{\sim}{T}$; and provided $|t_i| < 1$ (which we henceforth assume), observe that $g(\underset{\sim}{u};\underset{\sim}{t}) \geq 0$ for all $\underset{\sim}{u}$ and that $g(\underset{\sim}{u};\underset{\sim}{t}) > 1$ for $\underset{\sim}{u} \in A^c$. These assertions follow, the first from the positive definite character of $\underset{\sim}{G}(\underset{\sim}{t})$, the second on completing the square as $g(\underset{\sim}{u};\underset{\sim}{t}) = \underset{\sim}{u}_1'[\underset{\sim}{G}_{11}(\underset{\sim}{t})]^{-1}\underset{\sim}{u}_1 + S$, for example, S being a positive semidefinite quadratic form. We accordingly have

$$E[g(\underset{\sim}{U};\underset{\sim}{t})] = \int_{A\cup A^c} g(\underset{\sim}{u};\underset{\sim}{t})dF(\underset{\sim}{u}) \geq \int_{A^c} g(\underset{\sim}{u};\underset{\sim}{t})dF(\underset{\sim}{u}) \geq \int_{A^c} dF(\underset{\sim}{u}) = 1-P(A)$$

and it remains to evaluate $E[g(\underset{\sim}{U};\underset{\sim}{t})] = \mathrm{tr}[\underset{\sim}{G}(\underset{\sim}{t})]^{-1}\underset{\sim}{\Omega}$, where $\underset{\sim}{\Omega} = [\underset{\sim}{\Omega}_{ij}] = V(\underset{\sim}{U})$, $\underset{\sim}{\Omega}_{11} = \underset{\sim}{I}_r/\delta_1$, $\underset{\sim}{\Omega}_{12} = \underset{\sim}{D}/\delta_1^{\frac{1}{2}}\delta_2^{\frac{1}{2}}$, and $\underset{\sim}{\Omega}_{22} = \underset{\sim}{I}_s/\delta_2$. A reduction using partitioned matrices yields

$$\begin{aligned} E[g(\underset{\sim}{U};\underset{\sim}{t})] = {} & \mathrm{tr}(\underset{\sim}{I}_r-\underset{\sim}{T}\underset{\sim}{T}')^{-1}/\delta_1 - \mathrm{tr}\ \underset{\sim}{T}(\underset{\sim}{I}_s-\underset{\sim}{T}'\underset{\sim}{T})^{-1}\underset{\sim}{D}'/\delta_1^{\frac{1}{2}}\delta_2^{\frac{1}{2}} - \mathrm{tr}\ \underset{\sim}{T}'(\underset{\sim}{I}_r-\underset{\sim}{T}\underset{\sim}{T}')^{-1}\underset{\sim}{D}/\delta_1^{\frac{1}{2}}\delta_2^{\frac{1}{2}} \\ & + \mathrm{tr}(\underset{\sim}{I}_s-\underset{\sim}{T}'\underset{\sim}{T})^{-1}/\delta_2 \end{aligned}$$

which, from the special structure of $\underset{\sim}{T}$ and $\underset{\sim}{D}$, becomes

$$E[g(\underset{\sim}{U};\underset{\sim}{t})] = \frac{(s-r)}{\delta_2} + \sum_{i=1}^{r} [(\delta_1^{-1}+\delta_2^{-1}) - 2\rho_i t_i/\delta_1^{\frac{1}{2}}\delta_2^{\frac{1}{2}}]/(1-t_i^2).$$

On varying $\underset{\sim}{t}$, we obtain the smallest bound by minimizing $E[g(\underset{\sim}{U};\underset{\sim}{t})]$ with respect to $\{t_1,\ldots,t_r\}$. Term-wise differentiation yields the solutions $t_i = \{(\delta_1+\delta_2) - [(\delta_1+\delta_2)^2 - 4\rho_i^2\delta_1\delta_2]^{\frac{1}{2}}\}/2\rho_i\delta_1^{\frac{1}{2}}\delta_2^{\frac{1}{2}}$ satisfying $|t_i| < 1$; when substituted into an earlier expression these give the required bound.

When $\delta_1 = \delta_2 = \delta$ the bound simplifies as follows.

COROLLARY 4.1. For each $\delta > 0$ we have

$$P((\underset{\sim}{Y}_1-\underset{\sim}{\mu}_1)'\underset{\sim}{\Sigma}_{11}^{-1}(\underset{\sim}{Y}_1-\underset{\sim}{\mu}_1)\le\delta,(\underset{\sim}{Y}_2-\underset{\sim}{\mu}_2)'\underset{\sim}{\Sigma}_{22}^{-1}(\underset{\sim}{Y}_2-\underset{\sim}{\mu}_2)\le\delta) \ge 1-[s+\sum_{i=1}^{r}(1-\rho_i^2)^{\frac{1}{2}}]/\delta.$$

It may be noted from the proof that Theorem 4 provides the best bounds available depending only on dispersion parameters. Theorem 4 is stronger than the corresponding version of Theorem 1, yielding the latter when $\rho_1 = \ldots = \rho_r = 0$. When $r = s = 1$, Theorem 4 yields a result of Lal (1955) and Corollary 4.1 a result of Berge (1937) in terms of the simple correlation parameter ρ.

3. Applications. The foregoing inequalities apply in a variety of problems in statistics and applied probability. Some examples follow.

3.1 Ballistics. Given random impact coordinates having an arbitrary dispersion matrix, the probability of hitting an elliptical target may be rephrased in terms of the distribution of a definite quadratic form. Joint distributions of such forms arise in connection with salvos and multiple independently targeted reentry vehicles. In the case of a salvo of shots subject to a common aiming error, the required joint probabilities are determined by noncentral distributions of definite quadratic forms.

In particular, let $\{\underset{\sim}{y}_1,\ldots,\underset{\sim}{y}_k\}$ be the points of impact of k shots all aimed at the same point $\underset{\sim}{\xi}$ but intended for a target $A = \{\underset{\sim}{y}|(\underset{\sim}{y}-\underset{\sim}{\tau})'\underset{\sim}{B}(\underset{\sim}{y}-\underset{\sim}{\tau})\le\delta)$ with center at $\underset{\sim}{\tau}$. Suppose the dispersion matrices $V(\underset{\sim}{y}_i) = \underset{\sim}{\Sigma}$ are equal for $i=1,\ldots,k$, but no other constraints are imposed on the joint distribution of $\{\underset{\sim}{y}_1,\ldots,\underset{\sim}{y}_k\}$. A routine application of Corollary 1.2, with $\underset{\sim}{\mu}_i = \underset{\sim}{\xi}$ and $\underset{\sim}{\theta}_i = \underset{\sim}{\tau}$ for $i=1,\ldots,k$, yields the bound

$$P(E) \le k\gamma/\delta$$

for the event E that one or more shots miss the target, where $\gamma = \mathrm{tr}\underset{\sim}{B}\underset{\sim}{\Sigma}+(\underset{\sim}{\xi}-\underset{\sim}{\tau})'\underset{\sim}{B}(\underset{\sim}{\xi}-\underset{\sim}{\tau})$. Perhaps more useful is the probability that one or more hits are scored; from Theorem 2 comes the bound

$$P(S) \ge 1-\gamma/\delta$$

for the probability of the event S that one or more hits occur. We note that both bounds may be improved by centering the impact distribution on the target center, thereby eliminating the noncentrality parameter $(\underset{\sim}{\xi}-\underset{\sim}{\tau})'B(\underset{\sim}{\xi}-\underset{\sim}{\tau})$ from the bound and guaranteeing a smaller value for P(E) and a larger value for P(S). Whereas probabilistic analyses of ballistics systems usually are carried out under Gaussian assumptions (cf. Eckler (1969), for example), the foregoing developments apply to all joint impact distributions having moments of second order.

3.2 Statistical Quality Control. The variability of an industrial production process often is monitored using S^2 charts based on underlying Gaussian distributions. Let $\{\underset{\sim}{Y}_1,\underset{\sim}{Y}_2,...\}$ be vectors of n observations drawn on successive occasions; let $\{S_1^2,S_2^2,...\}$ be the corresponding sample variances; and let σ_0^2 be the control variance. The S^2 chart is a graph of $\{S_1^2,S_2^2,...\}$ against time on the horizontal scale; the process is asserted to be out of control and corrective action is taken whenever S_i^2/σ_0^2 exceeds the control limit c_α. Operating characteristics are determined by the distribution of the run length, i.e. the number, N, of successive samples drawn before the chart signals that the process is out of control. When $\{\underset{\sim}{Y}_1,\underset{\sim}{Y}_2,...\}$ are independent

and identically distributed, whatever the underlying distribution, the run-length distribution is the geometric distribution $G(t;\alpha)$ with parameter $\alpha = P(S_i^2/\sigma_0^2 > c_\alpha)$.

On choosing $\underset{\sim}{\theta} = \underset{\sim}{\mu}$ and $\underset{\sim}{B} = \underset{\sim}{I}_n - n^{-1}\underset{\sim}{1}_n\underset{\sim}{1}_n'$, where $\underset{\sim}{1}_n' = [1,\ldots,1]$ is n-dimensional, we infer from Corollary 1.1 that

$$P(S_i^2/\sigma_0^2 > c_\alpha) \le 1/c_\alpha$$

when the process is in control. Moreover, because the family $\{G(\cdot;\alpha);\ \alpha\epsilon(0,1)\}$ of geometric distributions is stochastically decreasing in α, we get the universal stochastic lower bound

$$P(N>t) \ge 1-G(t;c_\alpha^{-1})$$

for every run-length distribution. Similar conclusions apply when the process is not in control and $\sigma^2 > \sigma_0^2$.

3.3 Weak Laws of Large Numbers. Let $\{\underset{\sim}{Y}_1,\underset{\sim}{Y}_2,\ldots\}$ be n-dimensional outcomes in a sequence of experiments having homoscedastic errors $\{\underset{\sim}{\varepsilon}_1,\underset{\sim}{\varepsilon}_2,\ldots\}$ uncorrelated on successive occasions. We study consistency properties of the Gauss-Markov estimator $\tilde{\underset{\sim}{\beta}}_i = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}_i$ under the assumed model $\underset{\sim}{Y}_i = \underset{\sim}{X}\underset{\sim}{\beta}+\underset{\sim}{\varepsilon}_i$, given that the actual model is

$$\underset{\sim}{Y}_i = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{Z}_i\underset{\sim}{\gamma} + \varepsilon_i \ .$$

Let $\hat{\underset{\sim}{\beta}}_N = N^{-1}(\tilde{\underset{\sim}{\beta}}_1+\ldots+\tilde{\underset{\sim}{\beta}}_N)$; observe that $E(\hat{\underset{\sim}{\beta}}_N-\beta) = N^{-1}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\Sigma_1^N\underset{\sim}{Z}_i\underset{\sim}{\gamma}$; and, using the uncorrelatedness of successive error vectors, compute the dispersion matrix $V(\hat{\underset{\sim}{\beta}}_N) = N^{-1}\sigma^2\underset{\sim}{M}$ with $\underset{\sim}{M} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}$. From Corollary 1.1 we have

$$P(||\hat{\beta}_N-\beta||>\varepsilon) \le \frac{\sigma^2 \mathrm{tr} M}{N\varepsilon^2} + \frac{\gamma' G_N' G_N \gamma}{\varepsilon^2}$$

where $||\ ||$ is the Euclidean norm and $G_N = N^{-1}(X'X)^{-1}X' \sum_{i=1}^{N} Z_i$. If we now suppose that $\lim_{N\to\infty} G_N = 0$, i.e., $\lim_{N\to\infty} X'(N^{-1}\sum_{i=1}^{N} Z_i) = 0$, we find on taking limits that $\hat{\beta}_N$ is weakly consistent for β. In particular, if $Z_i = Z$ and $X'Z = 0$, then $\hat{\beta}_N$ is clearly consistent. Note that this approach avoids the usual assumption that $\{Y_1, Y_2, \ldots\}$ be mutually independent and that the model be correct.

3.4 Simultaneous Confidence Bounds. In a random sample of n observations from some scalar distribution having the mean μ and variance σ^2, let $\bar{Y}$ and S^2 respectively be the sample mean and variance. On choosing $B_1 = (n\sigma^2)^{-2} 1_n 1_n'$ and $B_2 = [(n-1)\sigma^2]^{-1}(I_n - n^{-1} 1_n 1_n')$, we apply Theorem 1 directly to obtain the bounds

$$P(|\bar{Y}-\theta| \le \sigma\delta_1, S^2/\sigma^2 \le \delta_2^2) \ge 1-[\frac{1}{n\delta_1^2} + \frac{(\mu-\theta)^2}{\sigma^2\delta_1^2} + \frac{1}{\delta_2^2}] \; .$$

When inverted using $\theta = \mu$ this gives a bound on the joint confidence coefficient for a region yielding two-sided limits for μ and a one-sided limit for σ^2. We note that no further improvement of this bound is available from Theorem 4, owing to the fact that $n^{\frac{1}{2}}\bar{Y}$ is uncorrelated with each of the residuals $\{(Y_1-\bar{Y}), \ldots, (Y_n-\bar{Y})\}$.

REFERENCES

Anderson, T.W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proc. Amer. Math. Soc. 6, 170-176.

Berge, P.O. (1937). A note on a form of Tchebycheff's theorem for two variables. Biometrika 29, 405-406.

Eckler, A.R. (1969). A survey of coverage problems associated with point and area targets. Technometrics 11, 561-589.

Hotelling, H. (1936). Relations between two sets of variates. Biometrika 28, 321-377.

Jensen, D.R. and Foutz, R.V. (1978). Markov inequalities on partially ordered spaces. Research Report No. S-6, Department of Statistics, VPI&SU, Blacksburg, Virginia, 24061.

Lal, D.N. (1955). A note on a form of Tchebycheff's inequality for two or more variables. Sankhyā 15, 317-320.

Wilks, S.S. (1962). Mathematical Statistics. John Wiley and Sons, New York.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Q-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) INEQUALITIES FOR JOINT DISTRIBUTIONS OF QUADRATIC FORMS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER Q-1
7. AUTHOR(s) D.R. Jensen		8. CONTRACT OR GRANT NUMBER(s) DAAG-29-78-G-0172
9. PERFORMING ORGANIZATION NAME AND ADDRESS Virginia Polytechnic Institute Blacksburg, Virginia 24061		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE March 1980
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE NA

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multidimensional Chebychev and Markov inequalities, sharpened Chebychev bounds, weak laws of large numbers, ballistic probabilities, statistical quality control.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Chebychev inequalities are given for joint central and noncentral distributions of k quadratic forms; these are sharpened when k = 2 using the canonical correlations of Hotelling. Complementary inequalities are found as versions of Markov's inequality. Applications are noted in ballistics, in statistical quality control, in establishing consistency of Gauss-Markov

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.

estimates under dependence, and in constructing conservative joint confidence sets depending on the underlying distribution only through its low order moments.